



Munich Personal RePEc Archive

Forecasting transaction counts with integer-valued GARCH models

Aknouche, Abdelhakim and Almohaimeed, Bader and Dimitrakopoulos, Stefanos

USTHB and Qassim University, Qassim University, Leeds University

11 July 2020

Online at <https://mpra.ub.uni-muenchen.de/101779/>

MPRA Paper No. 101779, posted 15 Jul 2020 09:18 UTC

Forecasting transaction counts with integer-valued *GARCH* models

Abdelhakim Aknouche^{*}, Bader Almohaimeed^{**}, and Stefanos

Dimitrakopoulos^{1***}

^{*}Department of Mathematics, College of Science, Qassim University (Saudi Arabia) & Faculty of Mathematics, University of Science and Technology Houari Boumediene (Algeria)

^{**}Department of Mathematics, College of Science, Qassim University, Saudi Arabia.

^{***}Division of Economics, Leeds University, UK

Abstract

Using numerous transaction data on the number of stock trades, we conduct a forecasting exercise with INGARCH models, governed by various conditional distributions. The model parameters are estimated with efficient Markov Chain Monte Carlo methods, while forecast evaluation is done by calculating point and density forecasts.

Keywords: Count time series, INGARCH models, MCMC

JEL CODE: C5, C22, G12

¹Correspondence to: Stefanos Dimitrakopoulos, s.dimitrakopoulos@leeds.ac.uk. All computations have been performed in the Advanced Research Computing (ARC) environment at Leeds University.

1 Introduction

In recent years, there has been a surge of interest in integer-valued generalized autoregressive conditional heteroscedastic (*INGARCH*) models (Fokianos et al., 2009; Doukhan et al., 2012; Christou and Fokianos, 2014; Chen et al., 2016; Davis and Liu, 2016; Ahmad and Francq, 2016; Aknouche et al., 2018). Such processes are designed to model integer-valued series that are characterized mainly by small values and overdispersion that can not be adequately accounted for by standard real-valued ARMA models; see also Cameron and Trivedi (2013).

In its original formulation (Grunwald et al., 2000; Rydberg and Shepard, 2000; Heinen, 2003), the *INGARCH* process had a Poisson conditional distribution with a time-varying intensity that was a linear function of its q lagged values and its p recent observations. Later, many generalizations of the Poisson *INGARCH* (*P-INGARCH*) were put forward that differed in their conditional distributions (Poisson, negative binomial, double Poisson, etc.) and/or their specifications for the conditional mean equation (linear, exponential, threshold); see, for example, Fokianos et al., (2009).

In spite of the large number of *INGARCH* models that have been proposed, the relevant literature seems to lack a unified forecasting comparison exercise, especially in a Bayesian framework. Therefore, using numerous empirical time series on the trade intensity of stocks, we evaluate the out-of-sample forecasting performance of various *INGARCH* models. Our set of competing *INGARCH* models includes those with the most popular conditional distributions, namely the Poisson, the (linear and quadratic) negative binomial and the double Poisson.

We estimate the model parameters by efficient Bayesian methods, in particular Markov Chain Monte Carlo (MCMC). The dispersion parameters are updated using an extremely efficient universal self-tuned sampler within Gibbs sampler, proposed by Martino et al., (2015), whilst for the *GARCH* parameters, the adaptive Metropolis adjusted Langevin (MALA) algorithm of Atchadé (2006) is exploited. The forecasting performance of the models is evaluated by calculating point and density forecasts.

The structure of the paper is as follows. Section 2 describes the models in question and Section 3 describes the calculation of point and density forecasts. Section 4 presents the empirical results. Section 5 concludes. An Online Appendix accompanies this paper.

2 INGARCH specifications

A stochastic process $\{Y_t, t \in \mathbb{Z}\}$ is said to be an INGARCH(p, q) if its conditional distribution is given by

$$Y_t | \mathcal{F}_{t-1} \sim f_{\lambda_t} \quad (1)$$

and

$$\lambda_t = \omega + \sum_{i=1}^q \alpha_i Y_{t-i} + \sum_{j=1}^p \beta_j \lambda_{t-j}, \quad (2)$$

where $\omega > 0, \alpha_i \geq 0$ and $\beta_j \geq 0$, \mathcal{F}_t is the σ -Algebra generated by $\{Y_{t-k}, k \geq 0\}$ and $f_{\lambda_t}(y_t) := f_{Y_t}(y_t | \mathcal{F}_{t-1})$ is a discrete distribution with mean λ_t .

In this paper we consider four distributions for $Y_t | \mathcal{F}_{t-1}$:

- The Poisson (P-INGARCH) (Heinen, 2003; Ferland et al., 2006); $Y_t | \mathcal{F}_{t-1} \sim \mathcal{P}(\lambda_t)$.
- The double Poisson (DP-INGARCH) (Heinen, 2003); $Y_t | \mathcal{F}_{t-1} \sim \mathcal{DP}(\lambda_t, \gamma)$, with $\gamma > 0$.
- The Negative binomial II (NB2-INGARCH) (Zhu, 2011, Christou and Fokianos, 2014; Davis and Liu, 2016); $Y_t | \mathcal{F}_{t-1} \sim \mathcal{NB}\left(r_2, \frac{r_2}{r_2 + \lambda_t}\right)$, with $r_2 > 0$.
- The Negative binomial I (NB1-INGARCH) (Aknouche and Francq, 2020); $Y_t | \mathcal{F}_{t-1} \sim \mathcal{NB}\left(r_1 \lambda_t, \frac{r_1}{r_1 + 1}\right)$, with $r_1 > 0$.

The functional forms of these distributions along with their conditional means and conditional variances are given in Table 1. The parameters γ, τ, r_1 and r_2 are usually called the dispersion parameters. As can be seen from Table 1, the conditional variance is linear in the intensity parameter λ_t for the Poisson, and NB1 cases, is approximately linear for the DP case and quadratic for the NB2 case. Under

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1, \quad (3)$$

the five models are stationary and ergodic with finite mean (Aknouche and Francq, 2020).

A brief description of the MCMC algorithms is given in the Online Appendix along with simulation studies.

3 Point and density forecasts

We conduct a recursive out-of-sample forecasting exercise in order to evaluate the predictive performance of the competing models. To this end, we compute point and density forecasts.

The conditional predictive density of the s -step ahead y_{t+s} , given the data $Y_t = (y_1, \dots, y_t)$ is given by

$$p(y_{t+s}|Y_t) = \int f(y_{t+s}|\Theta, Y_t)dp(\Theta|Y_t), \quad (4)$$

where Θ denotes the model parameters.

Using Monte Carlo integration, the above expression can be approximated by

$$\hat{p}(y_{t+s}|Y_t) = \frac{1}{R} \sum_{i=1}^R f(y_{t+s}|\Theta^{(i)}, Y_t), \quad (5)$$

where $\Theta^{(i)}$ is the posterior draw of Θ at iteration $i = 1, \dots, R$ (after the burn-in period).

The conditional predictive likelihood of y_{t+s} is the conditional predictive density of y_{t+s} evaluated at the observed y_{t+s}^o , namely, $p(y_{t+s} = y_{t+s}^o|Y_t)$. A usual metric for the evaluation of the density forecasts is the log predictive score (*LPS*) (Geweke and Amisano, 2011)

$$LPS = \sum_{t=t_0}^{T-s} \log p(y_{t+s} = y_{t+s}^o|Y_t), \quad (6)$$

where $t = t_0 + 1, \dots, T - s$ is the evaluation period. The higher the LPS value, the better the (out-of-sample) forecasting power of the model.

We also calculated s -step point forecasts. A usual metric for the evaluation of point forecasts is the root mean squared forecast error (RMSFE)

$$RMSFE = \sqrt{\frac{\sum_{t=t_0}^{T-s} (y_{t+s}^o - E(y_{t+s}|Y_t))^2}{T - s - t_0 + 1}}. \quad (7)$$

The lower the RMSFE value, the better the (out-of-sample) forecasting power of the model. In our analysis, $s = 1, 4$ and 8 .

4 Empirical analysis

4.1 Data

Our empirical data consist of four time series that record the number of trades for four stocks (Glatfelter Company (GLT), Wausau Paper Corporation (WPP), Empire District Electric Company (EDE), Ericsson B). For the first three stocks (GLT, WPP, EDE) we monitor the number of stock transactions in five-minute intervals between 9:45 AM and 4:00 PM. Each of these three series has $T = 2925$ observations and the time period is from January 3, 2005 to February 18, 2005. For the last stock (Ericsson B) the time series is of length $T = 460$ and records the number of transactions per minute between 9:35 AM and 17:14 PM on 2 July 2002. Plots of the time series and histograms are given in Figures 1 and 2. The data are strongly overdispersed. The estimation results are presented in the Online Appendix.

4.2 Forecasting results

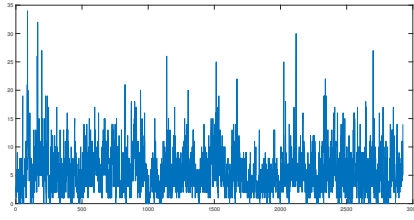
For our out-of-sample forecasting exercise, the evaluation period consists of the last 100 data points. The summary of the forecasting results is presented in Tables 2 (density forecasts) and 3 (point forecasts). The detailed forecasting results are reported in the Online Appendix. From Table 2 we can see that the NB2-INGARCH model is dominant for the first three data sets, producing better density forecasts than completing INGARCH specifications across all forecast horizons. The NB1-INGARCH yielded the best forecasting results for the Ericsson B data set only. The third-best model is the DP-INGARCH (see Online Appendix).

From Table 3, the results indicate that both the NB1-INGARCH and NB2-INGARCH models produce better point forecasts than the P-INGARCH and the DP-INGARCH models. In most of the cases (Online Appendix), the DP-INGARCH did better than the P-INGARCH.

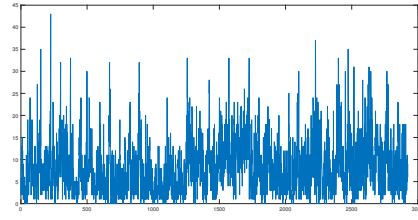
5 Conclusions

We conducted a Bayesian forecasting exercise using INGARCH models with various conditional distributions. Our empirical application concerned the number of stock trades. We found that the NB2-INGARCH model is superior, in terms of density forecasts, to other competing models

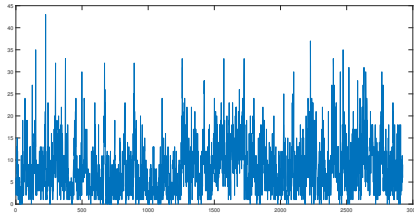
in predicting transaction counts, whereas the NB1-INGARCH and NB2-INGARCH models seem to dominate in point forecasting.



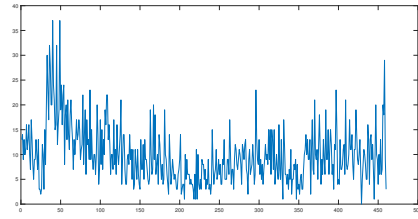
(a) GLT.



(b) EDE.

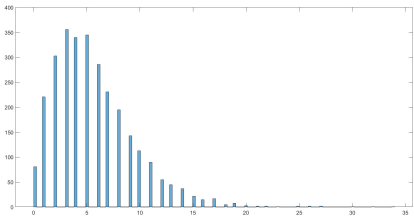


(c) WPP.

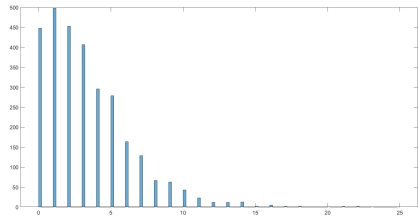


(d) Ericsson B

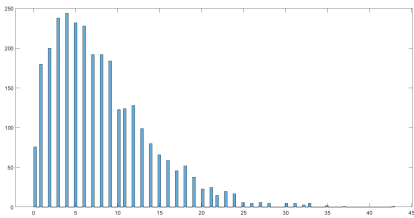
Figure 1: Empirical results: Time series plots for the four financial series.



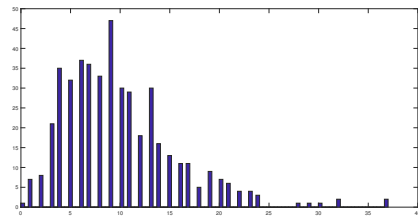
(a) GLT.



(b) EDE.



(c) WPP.



(d) Ericsson B

Figure 2: Empirical results: Histograms for the four financial time series.

Table 1: Various conditional distributions for the INGARCH model.

	Notation	$f_{Y_t}(y_t/\mathcal{F}_{t-1})$	$E(y_t \mathcal{F}_{t-1})$	$Var(y_t \mathcal{F}_{t-1})$
P	$\mathcal{P}(\lambda_t)$	$e^{-\lambda_t} \frac{\lambda_t^{y_t}}{y_t!}$	λ_t	λ_t
DP	$\mathcal{DP}(\lambda_t, \gamma)$	$\gamma^{1/2} e^{-\gamma\lambda_t} \frac{e^{-y_t} y_t^{y_t}}{y_t!} \left(\frac{e\lambda_t}{y_t}\right)^{\gamma y_t}$	λ_t	$\simeq \frac{1}{\gamma} \lambda_t$
NB1	$\mathcal{NB}\left(r_1\lambda_t, \frac{r_1}{r_1+1}\right)$	$\frac{\Gamma(y_t+r_1\lambda_t)}{y_t!\Gamma(r_1\lambda_t)} \left(\frac{r_1}{r_1+1}\right)^{r_1\lambda_t} \left(\frac{1}{r_1+1}\right)^{y_t}$	λ_t	$\left(1 + \frac{1}{r_1}\right) \lambda_t$
NB2	$\mathcal{NB}\left(r_2, \frac{r_2}{r_2+\lambda_t}\right)$	$\frac{\Gamma(y_t+r_2)}{y_t!\Gamma(r_2)} \left(\frac{r_2}{r_2+\lambda_t}\right)^{r_2} \left(\frac{\lambda_t}{r_2+\lambda_t}\right)^{y_t}$	λ_t	$\lambda_t + \frac{1}{r_2} \lambda_t^2$

Table 2: Summary table for the LPS results.

Data	$s = 1$	$s = 4$	$s = 8$
GLT	NB2-INGARCH	NB2-INGARCH	NB2-INGARCH
WPP	NB2-INGARCH	NB2-INGARCH	NB2-INGARCH
EDE	NB1-INGARCH	NB2-INGARCH	NB2-INGARCH
Ericsson B	NB1-INGARCH	NB1-INGARCH	NB1-INGARCH

Table 3: Summary table for the RMSFE results.

Data	$s = 1$	$s = 4$	$s = 8$
GLT	NB2-INGARCH	NB2-INGARCH	NB2-INGARCH
WPP	P-INGARCH	NB1-INGARCH	NB1-INGARCH
EDE	NB1-INGARCH	NB1-INGARCH	NB1-INGARCH
Ericsson B	NB2-INGARCH	NB2-INGARCH	NB2-INGARCH

References

- Ahmad, A., Francq, C. (2016). Poisson QMLE of count time series models. *Journal of Time Series Analysis* 37: 291-314.
- Aknouche, A., Bendjeddou, S., Touche, N. (2018). Negative binomial quasi-likelihood inference for general integer-valued time series models. *Journal of Time Series Analysis* 39: 192-211.
- Aknouche, A., Francq, C. (2020). Count and duration time series with equal conditional stochastic and mean orders. *Econometric Theory*, 1-33
- Atchadé, YF. (2006). An adaptive version for the Metropolis adjusted Langevin algorithm with a truncated drift. *Methodology and Computing in Applied Probability* 8: 235–254.
- Cameron, A., Trivedi, P. (2013). *Regression Analysis of Count Data*. Cambridge University Press.
- Chen, C.W.S., So, M., Li, J.C., Sriboonchitta, S. (2016). Autoregressive conditional negative binomial model applied to over-dispersed time series of counts. *Statistical Methodology* 31: 73–90.
- Christou, V., Fokianos, K. (2014). Quasi-likelihood inference for negative binomial time series models. *Journal of Time Series Analysis* 35: 55–78.
- Davis, R.A., Liu, H. (2016). Theory and inference for a class of observation-driven models with application to time series of counts. *Statistica Sinica* 26: 1673-1707.
- Doukhan, P., Fokianos K., Tjøstheim, D. (2012). On weak dependence conditions for Poisson autoregressions. *Statistics and Probability Letters* 82: 942–948.
- Ferland, R., Latour, A., Oraichi, D. (2006). Integer-valued GARCH process. *Journal of Time Series Analysis* 27: 923-942.
- Fokianos, K., Rahbek A., Tjøstheim, D. (2009). Poisson autoregression. *Journal of the American Statistical Association* 140: 1430–1439.
- Geweke, J., Amisano, G. (2011). Hierarchical Markov normal mixture models with applications to financial asset returns. *Journal of Applied Econometrics* 26:1–29.

Grunwald, G., Hyndman, R.J., Tedesco, L., Tweedie, R.L. (2000). Theory and methods: Non-Gaussian conditional linear AR(1) models. *Australian & New Zealand Journal of Statistics* 42: 479–495.

Heinen, A. (2003). Modelling time series count data: an autoregressive conditional Poisson model. *Available at SSRN 1117187*.

Martino, L., Yang, H., Luengo, D., Kannianen, J., Corande, J. (2015). A fast universal self-tuned sampler within Gibbs sampling. *Digital Signal Processing* 47: 68–83.

Rydberg, T.H., Shephard, N. (2000). BIN models for trade-by-trade data. Modelling the number of trades in a fixed interval of time. *Technical Report 0740, Econometric Society*.

Zhu, F. (2011). A negative binomial integer-valued GARCH model. *Journal of Time Series Analysis* 32: 54–67.

Online Appendix for: Forecasting transaction counts with integer-valued *GARCH* models

Abdelhakim Aknouche^{*}, Bader Almohaimeed^{**}, and Stefanos Dimitrakopoulos^{1***}

^{*}Department of Mathematics, College of Science, Qassim University (Saudi Arabia) & Faculty of Mathematics, University of Science and Technology Houari Boumediene (Algeria)

^{**}Department of Mathematics, College of Science, Qassim University, Saudi Arabia.

^{***}Division of Economics, Leeds University, UK

1 MCMC

We want to sample iteratively from the full conditional posteriors $\pi(\Delta|disp, y)$ and $\pi(disp|\Delta, y)$, where $\Delta = (\omega, \alpha, \beta)'$ and $disp$ represents the dispersion parameter, depending on the model. For Δ we used a truncated log-normal prior

$$\log(\Delta) \sim N(\mu_{\Delta}, \Sigma_{\Delta})1_{(\alpha+\beta<1)},$$

that satisfies the stationarity condition that $\alpha + \beta < 1$, whereas for the dispersion parameter we use a gamma prior

$$G(k_{disp}, m_{disp}).$$

Both conditionals are intractable and therefore we use Metropolis-Hastings type algorithms. For the update of the dispersion parameter we use the Fast Universal Self-Tuned Sampler (FUSS) of Martino et al., (2015)¹. It can be used to sample efficiently from univariate distributions. It consists of four steps. In the first step, an initial set of support points of the target distribution is chosen. In the second step, unused support points drop according to some pre-defined pruning criterion (for example, optimal minimax pruning strategy). In the third step, we have the construction of the independent proposal density, tailored to the shape of the target, with some appropriate pre-defined mechanism (for example interpolation). In the final step, a Metropolis-Hastings (MH) method is used.

For the update of Δ we use the adaptive MALA of Atchadé (2006) with a truncated drift. Defining the drift as

$$\forall X_n, D(X_n) = \frac{\delta}{\max(\delta, \|\nabla \log \pi(x_n)\|)} \nabla \log \pi(X_n)$$

¹Correspondence to: Stefanos Dimitrakopoulos, s.dimitrakopoulos@leeds.ac.uk. All computations have been performed in the Advanced Research Computing (ARC) environment at Leeds University.

¹The FUSS algorithm has better mixing properties than alternative MCMC methods such as slice sampling, MALA sampling, and Hamiltonian Monte Carlo sampling and is faster. The FUSS matlab function is accessible from Martino's webpage.

Algorithm: Adaptive MALA

Start with an initial point $X_0 \in \mathcal{U}$, a vector $\mu_0 \in \mathcal{U}$, a positive definite matrix Γ_0 , $\varepsilon > 0$, a sequence of positive step sizes $(\gamma_n)_{n \geq 1}$, $\bar{\tau} \in]0, 1[$ and $\sigma_0 > 0$.

Given the current value X_n and $(\mu_n, \Gamma_n, \sigma_n)$, let $\Lambda_n = \Gamma_n + \varepsilon \mathbf{I}_d$.

Generate $Y_{n+1} \sim \mathcal{N}\left(X_n + \frac{\sigma_n^2}{2} \Lambda_n D(X_n), \sigma_n^2 \Lambda_n\right)$ and generate $U \sim \mathcal{U}([0, 1])$.

Define $\alpha(X_n, Y_{n+1}) = \min\left(1, \frac{\pi(Y_{n+1})q_{\sigma_n, \Lambda_n}(X_n | Y_{n+1})}{\pi(X_n)q_{\sigma_n, \Lambda_n}(Y_{n+1} | X_n)}\right)$. If $U \leq \alpha(X_n, Y_{n+1})$, set $X_{n+1} = Y_{n+1}$.

Otherwise, set $X_{n+1} = X_n$.

Set :

$$\begin{aligned}\mu_{n+1} &= \mu_n + \gamma_n (X_{n+1} - \mu_n) \\ \Gamma_{n+1} &= \Gamma_n + \gamma_n ((X_{n+1} - \mu_n)(X_{n+1} - \mu_n)^\top - \Gamma_n) \\ \sigma_{n+1} &= \sigma_n + \gamma_n (\alpha(X_n, Y_{n+1}) - \bar{\tau}).\end{aligned}$$

the general algorithm is described as above, where $\pi()$ represents the posterior density and q_{σ_n, Λ_n} is the proposal, which in our case is the normal satisfying the stationarity condition. $\bar{\tau}$ is set, practically to 0.5 to achieve an acceptance rate of 50% and for numerical stability we set $\varepsilon = 10^{-6}$ and $\delta = 1000$. The sequence $(\gamma)_n \in N^*$ is chosen such that $\forall n, \gamma_n > 0, \sum_n \gamma_n = +\infty$ and $\gamma_n = O(n^{-\xi})$ with $1/2 < \xi \leq 1$.

2 Monte Carlo experiments

To assess the performance of the proposed Bayesian methodology we simulated various *INGARCH* series. Throughout our simulations, we generated $n=500$ and $n=1000$ data points from all models with various sets of real values of the parameters. These sample sizes are similar to those used in the empirical study. We run the samplers for 10000 iterations after discarding the initial 10000 cycles (burn-in period).

For the INGARCH parameter $\Delta = (\omega, \alpha, \beta)'$ we used a truncated log-normal prior

$$\log(\Delta) \sim N(\mu_\Delta, \Sigma_\Delta) 1_{(\alpha+\beta < 1)},$$

where $\mu_\Delta = (1, \log(0.1), \log(0.8))'$ and $\Sigma_\Delta = \text{diag}(10, 1, 1)$. For the dispersion parameters, we used the following gamma prior

$$G(5, 0.1).$$

To monitor the performance of our sampling algorithms, we estimated the inefficiency factor (IF); see Chib (2001). To monitor any lack of convergence, we also computed the Convergence Diagnostics (CD) statistic of Geweke (1992).

2.1 Simulation results for the P-INGARCH model

Table 1: Simulated data for the P-INGARCH (T=500)

True values	Mean	Stdev	IF	CD
$\omega = 1$	1.018	0.193	12.865	-0.708
$\alpha = 0.7$	0.641	0.041	15.541	-1.106
$\beta = 0.2$	0.248	0.046	16.795	1.026

Table 2: Simulated data for the P-INGARCH (T=1500)

True values	Mean	Stdev	IF	CD
$\omega = 1$	1.145	0.157	13.694	-1.389
$\alpha = 0.7$	0.671	0.028	13.265	-1.94
$\beta = 0.2$	0.207	0.033	13.471	1.949

Table 3: Simulated data for the P-INGARCH (T=500)

True values	Mean	Stdev	IF	CD
$\omega = 2$	2.108	0.418	22.098	1.298
$\alpha = 0.3$	0.286	0.035	21.815	1.650
$\beta = 0.6$	0.612	0.044	24.141	-1.804

Table 4: Simulated data for the P-INGARCH (T=1500)

True values	Mean	Stdev	IF	CD
$\omega = 2$	2.333	0.382	19.613	-0.442
$\alpha = 0.3$	0.271	0.027	32.791	-0.387
$\beta = 0.6$	0.614	0.036	34.447	0.443

2.2 Simulation results for the NB1-INGARCH model

Table 5: Simulated data for the NB1-INGARCH (T=500)

True values	Mean	Stdev	IF	CD
$\omega = 0.1$	0.070	0.016	12.073	-0.687
$\alpha = 0.7$	0.664	0.068	12.626	-1.011
$\beta = 0.2$	0.225	0.061	12.15	1.375
$r_1 = 4$	3.147	1.596	1.208	1.005

Table 6: Simulated data for the NB1-INGARCH (T=1500)

True values	Mean	Stdev	IF	CD
$\omega = 0.1$	0.097	0.014	13.324	-0.280
$\alpha = 0.7$	0.693	0.047	11.469	-0.829
$\beta = 0.2$	0.183	0.043	12.429	-0.023
$r_1 = 4$	3.735	0.792	1.2687	-0.472

Table 7: Simulated data for the NB1-INGARCH (T=500)

True values	Mean	Stdev	IF	CD
$\omega = 1$	1.193	0.272	22.033	0.504
$\alpha = 0.3$	0.294	0.037	21.711	1.349
$\beta = 0.6$	0.549	0.054	29.378	-0.970
$r_1 = 8$	6.574	2.090	1.030	0.361

Table 8: Simulated data for the NB1-INGARCH (T=1500)

True values	Mean	Stdev	IF	CD
$\omega = 1$	1.292	0.215	16.301	-2.680
$\alpha = 0.3$	0.303	0.027	17.708	-2.263
$\beta = 0.6$	0.522	0.043	18.519	2.818
$r_1 = 8$	7.101	2.375	1.022	-0.147

2.3 Simulation results for the NB2-INGARCH model

Table 9: Simulated data for the NB2-INGARCH (T=500)

True values	Mean	Stdev	IF	CD
$\omega = 1$	0.970	0.189	13.018	-0.879
$\alpha = 0.7$	0.652	0.047	15.223	-1.039
$\beta = 0.2$	0.244	0.050	16.47	0.946
$r_2 = 8$	9.013	1.244	1	1.391

Table 10: Simulated data for the NB2-INGARCH (T=1500)

True values	Mean	Stdev	IF	CD
$\omega = 1$	1.276	0.161	11.897	-1.524
$\alpha = 0.7$	0.692	0.031	12.667	-1.624
$\beta = 0.2$	0.176	0.033	12.447	2.585
$r_2 = 8$	8.093	0.711	1	-0.906

Table 11: Simulated data for the NB2-INGARCH (T=500)

True values	Mean	Stdev	IF	CD
$\omega = 2$	1.959	0.414	14.726	-1.188
$\alpha = 0.4$	0.422	0.047	13.091	-1.556
$\beta = 0.4$	0.392	0.062	13.63	1.896
$r_2 = 4$	3.911	0.301	1.486	0.877

Table 12: Simulated data for the NB2-INGARCH (T=1500)

True values	Mean	Stdev	IF	CD
$\omega = 2$	2.050	0.318	15.3	-0.411
$\alpha = 0.4$	0.380	0.032	12.198	-1.050
$\beta = 0.4$	0.413	0.049	15.073	0.761
$r_2 = 4$	4.036	0.201	3.393	0.890

2.4 Simulation results for the DP-INGARCH model

Table 13: Simulated data for the DP-INGARCH (T=500)

True values	Mean	Stdev	IF	CD
$\omega = 4$	4.823	0.945	68.29	0.326
$\alpha = 0.2$	0.158	0.040	67.554	1.534
$\beta = 0.5$	0.343	0.105	88.009	-0.689
$\gamma = 1$	1.508	0.095	1.006	-0.678

Table 14: Simulated data for the DP-INGARCH (T=1500)

True values	Mean	Stdev	IF	CD
$\omega = 4$	3.994	0.681	79.012	-1.254
$\alpha = 0.2$	0.178	0.026	26.58	-0.866
$\beta = 0.5$	0.408	0.077	60.079	1.264
$\gamma = 1$	1.391	0.063	1.0026	-1.114

Table 15: Simulated data for the DP-INGARCH (T=500)

True values	Mean	Stdev	IF	CD
$\omega = 1$	1.116	0.180	13.548	-1.563
$\alpha = 0.6$	0.471	0.049	13.935	-1.740
$\beta = 0.2$	0.181	0.064	13.467	2.154
$\gamma = 0.2$	0.292	0.018	1	0.234

Table 16: Simulated data for the DP-INGARCH (T=1500)

True values	Mean	Stdev	IF	CD
$\omega = 1$	1.167	0.126	11.768	-1.294
$\alpha = 0.6$	0.541	0.034	12.64	-1.270
$\beta = 0.2$	0.133	0.041	11.994	1.598
$\gamma = 0.2$	0.271	0.012	1.670	-0.093

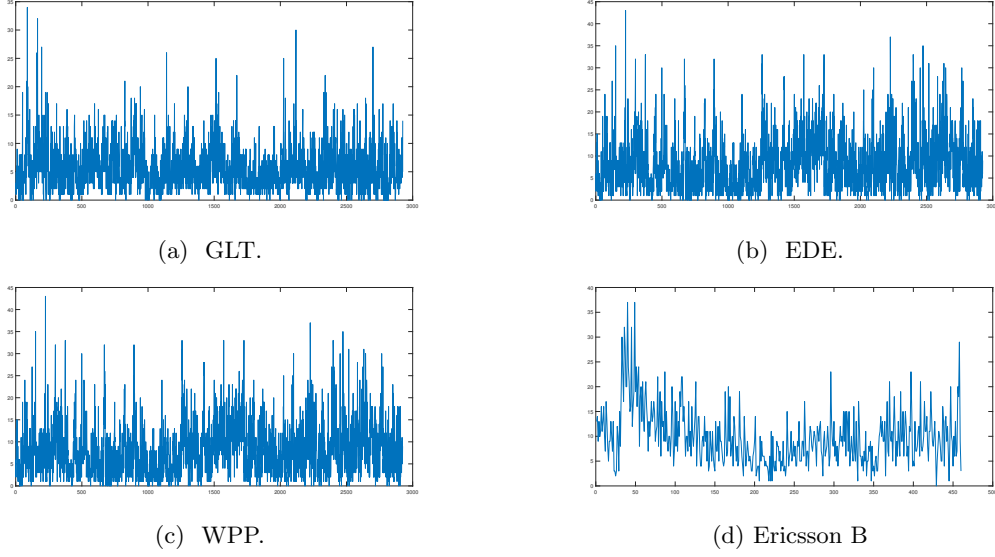


Figure 1: Empirical results: Time series plots for the four financial series.

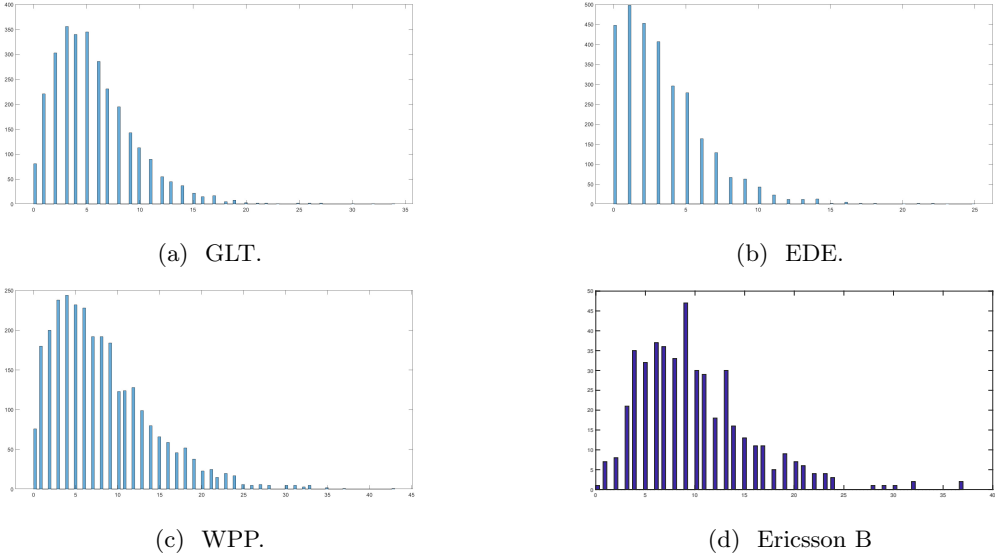


Figure 2: Empirical results: Histograms for the four financial time series.

3 Empirical analysis

3.1 Descriptive plots

It is given in Figures 1 and 2.

3.2 Estimation results

The hyperparameters for the prior distributions of the models in question are similar to those used in the simulation study. We run each algorithm for 5000 iterations after a burn-in period of 10000 cycles.

Table 17: Empirical results for P-INGARCH model

	GLT			EDE			WPP			Ericsson B		
	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD
ω	0.517 (0.052)	11.793	0.369	0.530 (0.052)	10.24	-0.165	0.859 (0.062)	9.515	-0.977	1.0957 (0.217)	24.324	-0.718
α	0.190 (0.010)	15.806	-1.365	0.224 (0.0117)	16.537	-0.071	0.270 (0.009)	11.985	-1.204	0.214 (0.021)	17.649	-0.999
β	0.718 (0.016)	13.634	0.344	0.615 (0.023)	11.385	-0.036	0.624 (0.014)	10.43	0.964	0.677 (0.036)	22.354	0.825

Standard deviation in parentheses. CD stands for Converge
Diagnostics and IF stands for Inefficiency Factor.

Table 18: Empirical results for NB1-INGARCH model

	GLT			EDE			WPP			Ericsson B		
	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD
ω	0.529 (0.084)	17.129	0.020	0.482 (0.070)	21.237	1.406	0.881 (0.118)	12.884	-1.551	0.708 (0.290)	32.748	1.315
α	0.180 (0.014)	18.272	-0.282	0.214 (0.017)	19.306	1.124	0.260 (0.017)	13.246	-0.304	0.194 (0.031)	29.208	0.305
β	0.724 (0.025)	18.053	0.014	0.639 (0.032)	23.909	-1.186	0.629 (0.027)	13.57	0.836	0.736 (0.050)	34.128	-0.637
r_1	0.950 (0.053)	1.181	3.068	0.747 (0.038)	1.895	0.419	0.467 (0.020)	1.815	2.856	0.706 (0.085)	1.296	-0.314

Standard deviation in parentheses. CD stands for Converge
Diagnostics and IF stands for Inefficiency Factor.

Table 19: Empirical results for NB2-INGARCH model

	GLT			EDE			WPP			Ericsson B		
	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD
ω	0.518 (0.081)	15.815	-0.625	0.519 (0.088)	20.959	-0.606	0.833 (0.111)	12.682	1.694	0.722 (0.321)	79.684	-0.319
α	0.193 (0.015)	18.281	0.433	0.224 (0.019)	15.846	-0.150	0.271 (0.017)	13.071	1.820	0.198 (0.030)	61.863	-1.165
β	0.713 (0.026)	18.408	0.027	0.618 (0.040)	19.732	0.554	0.623 (0.026)	12.053	-2.036	0.730 (0.049)	70.747	1.129
r_2	5.042 (0.201)	2.937	0.729	2.529 (0.208)	17.432	-2.338	3.565 (0.093)	294.78	5.597	6.504 (0.741)	1.038	-0.849

Standard deviation in parentheses. CD stands for Converge Diagnostics and IF stands for Inefficiency Factor.

Table 20: Empirical results for DP-INGARCH model

	GLT			EDE			WPP			Ericsson B		
	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD	Mean	IF	CD
ω	0.533 (0.093)	30.103	-0.266	0.536 (0.094)	19.439	-0.784	0.868 (0.123)	13.941	0.846	1.126 (0.343)	50.444	0.602
α	0.192 (0.016)	15.253	-0.531	0.225 (0.018)	21.537	0.009	0.270 (0.018)	17.535	0.229	0.215 (0.0347)	30.978	0.764
β	0.713 (0.028)	25.247	0.293	0.612 (0.041)	20.531	0.514	0.622 (0.029)	15.855	-0.634	0.673 (0.057)	53.553	-0.819
γ	0.467 (0.012)	1.927	0.667	0.427 (0.011)	1.712	-0.723	0.318 (0.008)	1.598	0.323	0.403 (0.026)	1.010	-0.286

Standard deviation in parentheses. CD stands for Converge
Diagnostics and IF stands for Inefficiency Factor.

3.3 Forecasting results

We report the ratio of the LPS value of the baseline model to that of a given model, with the baseline model being the P-INGARCH model. Hence, ratios greater than one indicate better density forecasts than the baseline model. Moreover, we subtract the RMSFEs value of a given model from that of the baseline model. So, positive values indicate better point forecasts.

We also calculated the Diebold and Mariano (1995) statistics, accounting also for the Harvey et al. (1997) finite-sample adjustment. The Diebold and Mariano (1995) approach is a test for equal predictive accuracy. Therefore, we tested whether the forecasting values (point and density forecasts) produced by the models are significantly different from those produced by the baseline model. The asterisk next to the reported density and point forecast value indicates that the respective model shows superior forecast performance relative to the baseline model.

3.3.1 Density forecasts

Table 21: LPS results (GLT).

Model	$s = 1$	$s = 4$	$s = 8$
P-INGARCH	1	1	1
NB1-INGARCH	0.8743	1.0871*	1.0833*
NB2-INGARCH	1.1118*	1.1110*	1.1091*
DP-INGARCH	1.0816*	1.0786*	1.0755*

Table 22: LPS results (WPP).

Model	$s = 1$	$s = 4$	$s = 8$
P-INGARCH	1	1	1
NB1-INGARCH	0.8744	1.0533*	1.0575*
NB2-INGARCH	1.0690*	1.0619*	1.0659*
DP-INGARCH	1.0611*	1.0489*	1.0520*

Table 23: LPS results (EDE).

Model	$s = 1$	$s = 4$	$s = 8$
P-INGARCH	1	1	1
NB1-INGARCH	1.4299*	1.1295*	1.1402*
NB2-INGARCH	1.1464*	1.1503*	1.1594*
DP-INGARCH	1.1104*	1.1136*	1.1254*

Table 24: LPS results (Ericsson B).

Model	$s = 1$	$s = 4$	$s = 8$
P-INGARCH	1	1	1
NB1-INGARCH	1.3312*	1.1469*	1.1854*
NB2-INGARCH	1.1294*	1.0254 *	1.169*
DP-INGARCH	1.2290*	1.0478*	1.0256*

3.3.2 Point forecasts

Table 25: RMSFE results (GLT).

Model	$s = 1$	$s = 4$	$s = 8$
P-INGARCH	0	0	0
NB1-INGARCH	-0.5831	-0.0058	-0.0035
NB2-INGARCH	0.0151*	0.0217*	0.0237*
DP-INGARCH	0.0033*	0.0015*	0.0110*

Table 26: RMSFE results (WPP).

Model	$s = 1$	$s = 4$	$s = 8$
P-INGARCH	0	0	0
NB1-INGARCH	-0.5740	0.0231*	0.0251*
NB2-INGARCH	-0.0056	-0.0075	-0.0024
DP-INGARCH	-0.0003	-0.0049	0.0029*

Table 27: RMSFE results (EDE).

Model	$s = 1$	$s = 4$	$s = 8$
P-INGARCH	0	0	0
NB1-INGARCH	0.0888*	0.0353*	0.0421*
NB2-INGARCH	0.0143*	0.0148*	0.0095*
DP-INGARCH	0.0001*	0.0017*	0.0024*

Table 28: RMSFE results (Ericsson B).

Model	$s = 1$	$s = 4$	$s = 8$
P-INGARCH	0	0	0
NB1-INGARCH	0.1254*	0.1058*	0.0856*
NB2-INGARCH	0.1467*	0.1364*	0.1743*
DP-INGARCH	0.1224*	0.0145*	0.1346*

References

- Atchadé, YF. (2006). An adaptive version for the Metropolis adjusted Langevin algorithm with a truncated drift. *Methodology and Computing in Applied Probability* 8: 235–254.
- Chib, S. (2001). Markov chain Monte Carlo methods: computation and inference. In: *Handbook of Econometrics*, Volume 5. North Holland, Amsterdam, pp. 3569-3649.
- Diebold, F., Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13: 253–263.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: *Bayesian Statistics 4*, Oxford: Clarendon Press, pp. 641-649.
- Harvey, D., Leybourne, S., Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13: 281–291.
- Martino, L., Yang, H., Luengo, D., Kannianen, J., Corande, J. (2015). A fast universal self-tuned sampler within Gibbs sampling. *Digital Signal Processing* 47: 68–83.